

MILLNVLRICI	IVCLVNDGAG	KHSEGRERTK	TYSLNSRGYF	40	
RKERGARRSK	ILLVNTKG	LD	EPHIGHGDFG	LVAELFDSTR	80
THTNRKEPDM	NKV	KLFSTVA	HGNKSARRKA	YNGSRRNIFS	120
RRSF	DKRNT	E VTEKPGAKMF	WNNFLVKMNG	APQNTSHGSK	160
AQEIMKEACK	TLPFTQNI	VH	ENCDRMVIQN	NLCFGKCISL	200
HVPNQQDRRN	TCSHCLPSKF	TLNH	TLNCT	GSKNVVKVVM	240
MVEECTCEAH	KSNF	HQTAQF	NMDT	STTLHH	270

Figure 1. Deduced amino acid sequence of *Xenopus cerberus* protein. SEQ ID NO:1.

Figure 2. Nucleotide sequence of the full-length cerberus DNA derived from the *Xenopus* organizer. The sense strand is on top (in the 5' to 3' direction) and the antisense strand on the bottom line (on the opposite direction). SEQ ID NO:2.

GAATTCCCAG CAAGTCGCTC AGAACACTG CAGGGCTAG ATATCATACA ATGTTACTAA	60
CTTAAGGGTC GTTCAGCGAG TCTTGTGAC GTCCCAGATC TATAGTATGT TACAATGATT	
ATGTACTCAG GATCTGTATT ATCGTCTGCC TTGTGAATGA TGGAGCAGGA AAACACTCAG	120
TACATGAGTC CTAGACATAA TAGCAGACGG AACACTTACT ACCTCGTCCT TTTGTGAGTC	
AAGGACGAGA AAGGACAAAA ACATATTAC TTAACAGCAG AGGTTACTTC AGAAAAGAAA	180
TTCCCTGCTCT TTCCCTGTTT TGTATAAGTG AATTGTCGTC TCCAATGAAG TCTTTCTTT	
GAGGAGCACG TAGGAGCAAG ATTCTGCTGG TGAATACTAA AGGTCTGAT GAACCCCACA	240
CTCCTCGTGC ATCCTCGTTC TAAGACGACC ACTTATGATT TCCAGAACTA CTTGGGGTGT	
TTGGGCATGG TGATTTTCGC TTAGTAGCTG AACTATTTGA TTCCACCAGA ACACATACAA	300
AACCCGTACC ACTAAAAGCG AATCATCGAC TTGATAAACT AAGGTGGTCT TGTGTATGTT	
ACAGAAAAGA GCCAGACATG AACAAAGTCA AGCTTTTCTC AACAGTGCC CATGGAAACA	360
TGTCTTTCTC CGGTCTGTAC TTGTTTCAGT TCGAAAAGAG TTGTCAACGG GTACCTTTGT	
AAAGTGCAAG AAGAAAAGCT TACAATGGTT CTAGAAGGAA TATTTTCCT CGCCGTTCTT	420
TTTCACGTTTC TTCTTTTCGA ATGTTACCAA GATCTTCCTT ATAAAAGGA GCGGCAAGAA	
TTGATAAAAG AAATACAGAG GTTACTGAAA AGCCTGGTGC CAAGATGTC TGGAAACAATT	480
AACTATTTTC TTTATGTCTC CAATGACTTT TCGGACCACG GTTCTACAAG ACCTTGTAA	
TTTGGTTAA AATGAATGGA GCCCCACAGA ATACAAGCCA TGGCAGTAAA GCACAGGAAA	540
AAAACCAATT TTACTTACCT CGGGGTGTCT TATGTTCGGT ACCGTCAATT CGTGTCTTT	
TAATGAAAGA AGCTTGCAAA ACCTTGTCTT TCACTCAGAA TATTGTACAT GAAAACGTG	600
ATTACTTTCT TCGAACGTTT TGGAACAAAA AGTGAGTCTT ATAACATGTA CTTTGACAC	
ACAGGATGGT GATACAGAAC AATCTGTGCT TTGGTAAATG CATCTCTC CATGTTCCAA	660
TGTCCTACCA CTATGTCTG TTAGACACGA ACCATTTAC GTAGAGAGAG GTACAAGGTT	
ATCAGCAAGA TCGACGAAAT ACTTGTCTCC ATTGCTTGCC GTCCAAATTT ACCCTGAACC	720
TAGTCGTTCT AGCTGCTTTA TGAACAAGGG TAACGAACGG CAGGTTAAA TGGGACTTGG	
ACCTGACGCT GAATTGTACT GGATCTAAGA ATGTAGTAA GGTTGTCATG ATGGTAGAGG	780
TGGACTGCGA CTTAACATGA CCTAGATTCT TACATCATT CCAACAGTAC TACCATCTCC	
AATGCACGTG TGAAGCTCAT AAGAGCAACT TCCACCAAC TGCACAGTTT AACATGGATA	840
TTACGTGCAC ACTTCGAGTA TTCTCGTGA AGGTGGTTG ACGTGTCAA TTGTACCTAT	
CATCTACTAC CCTGCACCAT TAAAGGACTG CCATACAGTA TGGAAATGCC CTTTTGTTGG	900
GTAGATGATG GGACGTGGTA ATTCCTGAC GGATGTGAT ACCTTTACGG GAAAACAACC	
AATATTTGTT ACATACTATG CATCTAAAGC ATTATGTTGC CTTCTATTTC ATATAACCAC	960
TTATAAACAA TGTATGATAC GTAGATTTG TAATACAACG GAAGATAAAG TATATTGGTG	
ATGGAATAAG GATTGTATGA ATTATAATTA ACAAAATGGCA TTTTGTGAA CATGCAAGAT	1020
TACCTTATTC CTAACATACT TAATATTAAT TGTTTACCGT AAAACACATT GTACGTTCTA	

CTCTGTTCCA TCAGTTGCAA GATAAAAGGC AATATTTGTT TGACTTTTT TCTACAAAAT 1080
GAGACAAGGT AGTCAACGTT CTATTTCCG TTATAAACAA ACTGAAAAAA AGATGTTTA
GAATACCCAA ATATATGATA AGATAATGGG GTCAAAACTG TTAAGGGGTA ATGTAATAAT 1140
CTTATGGGTT TATATACTAT TCTATTACCC CAGTTTGAC AATTCCCCAT TACATTATTA
AGGGACTAAG TTTGCCAGG AGCAGTGACC CATAACAACC AATCAGCAGG TATGATTTAC 1200
TCCCTGATTC AAACGGGTCC TCGTCACTGG GTATTGTTGG TTAGTCGTCC ATACTAAATG
TGGTCACCTG TTTAAAAGCA AACATCTTAT TGGTTGCTAT GGGTTACTGC TTCTGGGCAA 1260
ACCAGTGGAC AAATTTCGT TTGTAGAATA ACCAACGATA CCCAATGACG AAGACCCGTT
AATGTGTGCC TCATAGGGGG GTTAGTGTGT TGTGTACTGA ATAAATTGTA TTTATTTCAT 1320
TTACACACGG AGTATCCCC CAATCACACA ACACATGACT TATTTAACAT AAATAAAGTA
TGTTACAAA AAAAAAAA
ACAATGTTT TTTTTTT

Fig. 2. (Continuation page 2, SEQ ID NO:2).

MSRTRKVDSL LLLAIPGLAL LLLPNAYCAS CEPVRIPMCK SMPWNMTKMP NHLHHSTQAN 60
AILAIEQFEG LLTTECSQDL LFFILCAMYAP ICTIDFQHEP IKPCKSVCER ARAGCEPILI 120
KYRHTWPESL ACEELPVYDR GVCISPEAIV TVEQGTD SMP DFSMDSNNGN CGSGREHCKC 180
KPMKATQKTY LKNNNYVIR AKVKEVKVKC HDATAIVEVK EILKSSLVNI PKDTVTLYTN 240
SGCLCPQLVA NEEYIIMGYE DKERTRLLL V EGSLAEKWRD RLAKKVKRWD QKLRRPRKSK 300
DPVAPIPNKN SNSRQARS

Figure 3. Deduced amino acid sequence of Xenopus frazzled protein. SEQ ID NO:3.

Figure 4. Nucleotide sequence of the full-length frazzled cDNA derived from the *Xenopus* organizer. The sense strand of the DNA on top (5' to 3' direction) and the antisense strand on the bottom line (opposite direction). SEQ ID NO:4.

GAATTCCCTT TCACACAGGA CTCCTGGCAG AGGTGAATGG TTAGCCCTAT GGATTTGGTT	60
CTTAAGGGAA AGTGTGTCTT GAGGACCGTC TCCACTTACC AATCGGGATA CCTAAACCAA	
TGTGATTTT GACACATGAT TGATTGCTTT CAGATAGGAT TGAAGGACTT GGATTTTAT	120
ACAACAAAAA CTGTGTACTA ACTAACGAAA GTCTATCCTA ACTTCCTGAA CCTAAAAATA	
CTAATTCTGC ACTTTAAAT TATCTGAGTA ATTGTTCATT TTGTATTGGA TGGGACTAAA	180
GATTAAGACG TGAAAATTAA ATAGACTCAT TAACAAGTAA AACATAACCT ACCCTGATTT	
GATAAACTTA ACTCCTTGC TTTGACTTGC CCATAAACTA TAAGGTGGGG TGAGTTGTAG	240
CTATTTGAAT TGAGGAACGA AAACGTAAACG GGTATTTGAT ATTCCACCCCC ACTCAACATC	
TTGCTTTAC ATGTGCCAG ATTTCCCTG TATTCCCTGT ATTCCCTCTA AAGTAAGCCT	300
AACGAAAATG TACACGGGTC TAAAAGGGAC ATAAGGGACA TAAGGGAGAT TTCATTGGAA	
ACACATACAG GTTGGGCAGA ATAACAATGT CTCGAACAG GAAAGTGGAC TCATTACTGC	360
TGTGTATGTC CAACCCGTCT TATTGTTACA GAGCTTGTTC CTTTCACCTG AGTAATGACG	
TACTGGCCAT ACCTGGACTG GCGCTTCTCT TATTACCCAA TGCTTACTGT GCTTCGTGTG	420
ATGACCGGTA TGGACCTGAC CGCGAAGAGA ATAATGGGTT ACGAATGACA CGAACGACAC	
AGCCTGTGCG GATCCCCATG TGCAAATCTA TGCCATGGAA CATGACCAAG ATGCCCAACC	480
TGGACACCGC CTAGGGGTAC ACGTTTAGAT ACGGTACCTT GTACTGGTTC TACGGGTTGG	
ATCTCCACCA CAGCACTCAA GCCAATGCCA TCCTGGCAAT TGAACAGTTT GAAGGTTTGC	540
TAGAGGTGGT GTCGTGAGTT CGGTTACGGT AGGACCGTTA ACTTGTCAAA CTTCCAAACG	
TGACCACTGA ATGTAGCCAG GACCTTTGT TCTTCTGTG TGCCATGTAT GCCCCCATT	600
ACTGGTACT TACATCGGTC CTGGAAAACA AGAAAGACAC ACGGTACATA CGGGGGTAAA	
GTACCATCGA TTTCCAGCAT GAACCAATTAA AGCCTTGCAA GTCCGTGTGC GAAAGGGCCA	660
CATGGTAGCT AAAGGTGTA CTTGGTTAAT TCGGAACGTT CAGGCACACG CTTTCCCCGT	
GGGCCGGCTG TGAGCCCATT CTCATAAAAGT ACCGGCACAC TTGGCCAGAG AGCCTGGCAT	720
CCCGGCCGAC ACTCGGGTAA GAGTATTCA TGGCCGTGTG AACCGGTCTC TCGGACCGTA	
GTGAAGAGCT GCCCCATAT GACAGAGGAG TCTGCATCTC CCCAGAGGCT ATCGTCACAG	780
CACTTCTCGA CGGGCATATA CTGTCTCCTC AGACGTAGAG GGGTCTCCGA TAGCAGTGT	
TGGAACAAAG AACAGATTCA ATGCCAGACT TCTCCATGGA TTCAAAACAT GGAATTGCG	840
ACCTTGTCC TTGTCTAAGT TACGGTCTGA AGAGGTACCT AAGTTGTTA CCTTTAACGC	
GAAGCGGCAG GGAGCACTGT AAATGCAAGC CCATGAAGGC AACCCAAAAG ACGTATCTCA	900
CTTCGCCGTC CCTCGTGACA TTTACGTTCG GGTACTTCG TTGGGTTTC TGCATAGAGT	
AGAATAATTA CAATTATGTA ATCAGAGCAA AAGTGAAGA GGTGAAGTG AAATGCCACG	960
TCTTATTAAT GTTAATACAT TAGTCTCGTT TTCACTTCT CCACCTTCAC TTACGGTGC	
ACGCAACAGC AATTGTGGAA GTAAAGGAGA TTCTCAAGTC TTCCCTAGTG AACATTCCTA	1020
TGCGTTGTGCG TTAACACCTT CATTTCCTCT AAGAGTTCAAG AAGGGATCAC TTGTAAGGAT	

AAGACACAGT GACACTGTAC ACCAACTCAG GCTGCTTGTG CCCCCAGCTT GTTCCAATG TTCTGTCA CTGTGACATG TGGTTGAGTC CGACGAACAC GGGGGTCGAA CAACGGTTAC	1080
AGGAATACAT AATTATGGGC TATGAAGACA AAGAGCGTAC CAGGCTTCTA CTAGTGGAAAG TCCTTATGTA TTAATACCCG ATACTTCTGT TTCTCGCATG GTCCGAAGAT GATCACCTTC	1140
GATCCTTGGC CGAAAAATGG AGAGATCGTC TTGCTAAGAA AGTCAAGCGC TGGGATCAAA CTAGGAACCG GCTTTTACCC TCTCTAGCAG AACGATTCTT TCAGTTCGCG ACCCTAGTTT	1200
AGCTTCGACG TCCCAGGAAA AGCAAAGACC CCGTGGCTCC AATTCCCAAC AAAAACAGCA TCGAAGCTGC AGGGTCCTT TCGTTCTGG GGCAACGAGG TTAAGGGTTG TTTTGTGCGT	1260
ATTCCAGACA AGCGCGTAGT TAGACTAACG GAAAGGTGTA TGGAAACTCT ATGGACTTTG TAAGGTCTGT TCGCGCATCA ATCTGATTGC CTTTCCACAT ACCTTTGAGA TACCTGAAAC	1320
AAACTAAGAT TTGCATTGTT GGAAGAGCAA AAAAGAAATT GCACTACAGC ACGTTATATT TTGATTCTA AACGTAACAA CCTTCTCGTT TTTTCTTTAA CGTGATGTCG TGCAATATAA	1380
CTATTGTTTA CTACAAGAAG CTGGTTAGT TGATTGAGT TCTCCTTCC TTCTTTTTT GATAACAAAT GATGTTCTTC GACCAATCA ACTAACATCA AGAGGAAAGG AAGAAAAAAA	1440
TTATAACTAT ATTTGCACGT GTTCCCAGGC AATTGTTTA TTCAACTTCC AGTGACAGAG AATATTGATA TAAACGTGCA CAAGGGTCCG TTAACAAAAT AAGTTGAAGG TCACTGTCTC	1500
CAGTGACTGA ATGTCTCAGC CTAAGAAAGC TCAATTCAATT TCTGATCAAC TAATGGTGAC GTCACTGACT TACAGAGTCG GATTCTTCG AGTTAAGTAA AGACTAGTTG ATTACCACTG	1560
AAGTGTGTTGA TACTTGGGGA AAGTGAACTA ATTGCAATGG TAAATCAGAG AAAAGTTGAC TTCACAAACT ATGAACCCCT TTCACTTGAT TAACGTTACC ATTTAGTCTC TTTTCAACTG	1620
CAATGTTGCT TTTCTGTAG ATGAACAAAGT GAGAGATCAC ATTTAAATGA TGATCACTT GTTACAACGA AAAGGACATC TACTGTTCA CTCTCTAGTG TAAATTTACT ACTAGTGAAA	1680
CCATTTAATA CTTTCAGCAG TTTAGTTAG ATGACATGTA GGATGCACCT AAATCTAAAT GGTAAATTAT GAAAGTCGTC AAAATCAATC TACTGTACAT CCTACGTGGA TTTAGATTAA	1740
ATTTTATCAT AAATGAAGAG CTGGTTTAGA CTGTATGGTC ACTGTTGGGA AGGTAAATGC TAAATAGTA TTACTTCTC GACCAAAATCT GACATACCAAG TGACAACCCCT TCCATTTACG	1800
CTACTTTGTC AATTCTGTTT TAAAAATTGC CTAAAATAAT ATTAAGTCCT AAATAAAAAA GATGAAACAG TTAAGACAAA ATTTTAACG GATTTATTAA TAATTCAAGGA TTTATTTTTT	1860
AAAAAAAAAAA AAAAA TTTTTTTTTT TTTTT	

Fig. 4. (Continuation page 2, SEQ ID NO:4).

MLLFRAIPM	LLLGLMVLQT	DCEIAQYYID	EEEPPGTVIA	VLSQHSIFNT	TDIPATNFRL	60
MKQFNNSLIG	VRESDGQLSI	MERIDREQIC	RQSLHCNLAL	DVVSFSKGHF	KLLNVKVEVR	120
DINDHSPHFP	SEIMHVEVSE	SSVGTRIPL	EIAIDEDVGS	NSIQNFQISN	NSHFSIDVLT	180
RADGVKYADL	VLIMRELDREI	QPTYIMELLA	MDGGVPSLSG	TAVVNIRVLD	FNDNSPVFER	240
STIAVLDLVED	APLGYLLEL	HATDDDEGVN	GEIVYGFSTL	ASQEVRLFK	INSRTGSVTL	300
EGQVDFETKQ	TYEFEVQAQD	LGPNPLTATC	KVTVHILDVN	DNTPAITITP	LTtvNAGVAY	360
IPETATKENF	IALISTTDRA	SGSNGQVRCT	LYGHEHFKLQ	QAYEDSYMIV	TTSTLDRENI	420
AAYSLTVVAE	DLGFPSSLTK	KYYTVKVSDE	NDNAPVFSKP	QYEASILENN	APGSYITTVI	480
ARDSDSDQNG	KVNYRLVDAK	VMGQLTTFV	SLDADSGVLR	AVRSIDYEKL	KQLDFEIEAA	540
DNGIPQLSTR	VQLNLRIVDQ	NDNCPVITNP	LLNNNGSGEVL	LPISAPQNYL	VFQLKAEDSD	600
EGHNSQLFYT	ILRDPSRLFA	INKESGEVFL	KKQLNSDHSE	DLSIVVAVYD	LGRPSLSTNA	660
TVKFILTDSE	PSNVEVVIQ	PSAEEQHQID	MSIIFIAVLA	GGCAGLLLAI	FFVACTCKKK	720
AGEFKQVPEQ	HGTCNEERLL	STPSPQSVSS	SLSQSESQL	SINTESENCs	VSSNQEQQHQQ	780
TGIKHSISVP	SYHTSGWHLD	NCAMSIISGHS	HMGHISTKVQ	WAKEIVTSMT	VTLILVENQK	840
RRALSSQCRH	KPVLNTQMNQ	QGSMDMPITIS	ATESTRVQKM	GTAHCNMKRA	IDCLTL	

Figure 5. Deduced amino acid sequence of the Xenopus PAPC (paraxial protocadherin) protein. It encodes a member of the cadherin family of transmembrane proteins that has dorsalizing activity when constructs are injected into Xenopus embryos. SEQ ID NO:5.

Figure 6. Nucleotide sequence of the full-length PAPC cDNA derived from the Xenopus organizer. The sense strand of the DNA is shown in the top line (in the 5' to 3' direction), and the bottom line shows the antisense strand (opposite orientation). SEQ ID NO:6.

GAATTCCAG AGATGAACCTC CTTGAGATTG TTTTAAATGA CTGCAGGTCT GGAAGGATTCTTAAGGGTC TCTACTTGAG GAACTCTAAC AAAATTTACT GACGTCCAGA CCTTCCTAAC	60
ACATTGCCAC ACTGTTCTA GGCATGAAAA AACTGCAAGT TTCAACTTTG TTTTGGTCTG TGTAACGGTG TGACAAAGAT CCGTACTTT TTGACGTTCA AAGTTGAAAC AAAACCCACG	120
AACTTTGATT CTTCAAGATG CTGCTCTCT TCAGAGCCAT TCCAATGCTG CTGTTGGGAC TTGAAACTAA GAAGTTCTAC GACGAAGAGA AGTCTCGGTA AGGTTACGAC GACAACCTG	180
TGATGGTTTT ACAAAACAGAC TGTGAAATTG CCCAGTACTA CATAGATGAA GAAGAACCCCC ACTACCAAAA TGTTGTCTG ACACCTTAAC GGGTCATGAT GTATCTACTT CTTCTTGGGG	240
CTGGCACTGT AATTGCACTG TTGTCACAAC ACTCCATATT TAACACTACA GATATACCTG GACCGTGACA TTAACGTCAC AACAGTGTG TGAGGTATAA ATTGTGATGT CTATATGGAC	300
CAACCAATT CCGTCTAATG AAGCAATTAA ATAATTCCCT TATCGGAGTC CGTGAGAGTG GTTGGTTAAA GGCAGATTAC TTCGTTAAAT TATTAAGGGAA ATAGCCTCAG GCACTCTCAC	360
ATGGGCAGCT GAGCATCATG GAGAGGATTG ACCGGGAGCA AATCTGCAGG CAGTCCCTTC TACCCGTCGA CTCGTAGTAC CTCTCCTAAC TGGCCCTCGT TTAGACGTCC GTCAGGGAAAG	420
ACTGCAACCT GGCTTGGAT GTGGTCAGCT TTTCAAAGG ACACCTCAAG CTTCTGAACG TGACGTTGGA CCGAAACCTA CACCAGTCGA AAAGGTTCC TGTGAAGTTC GAAGACTTGC	480
TGAAAGTGGA GGTGAGAGAC ATTAATGACC ATAGCCCTCA CTTTCCCAAGT GAAATAATGC ACTTTCACCT CCACTCTCTG TAATTACTGG TATCGGGAGT GAAAGGGTCA CTTTATTACG	540
ATGTGGAGGT GTCTGAAAGT TCCTCTGTGG GCACCCAGGAT TCCTTAGAA ATTGCAATAG TACACCTCCA CAGACTTCA AGGAGACACC CGTGGTCCTA AGGAAATCTT TAACGTTATC	600
ATGAAGATGT TGGGTCCAAC TCCATCCAGA ACTTTCAGAT CTCAAATAAT AGCCACTTCA TACTTCTACA ACCCAGGTTG AGGTAGGTCT TGAAAGTCTA GAGTTTATTA TCGGTGAAGT	660
GCATTGATGT GCTAACCCAGA GCAGATGGGG TGAAATATGC AGATTTAGTC TTAATGAGAG CGTAACCTACA CGATTGGTCT CGTCTACCCC ACTTTATACG TCTAAATCAG AATTACTCTC	720
AACTGGACAG GGAAATCCAG CCAACATACA TAATGGAGCT ACTAGCAATG GATGGGGGTG TTGACCTGTC CCTTTAGGTC GGTTGTATGT ATTACCTCGA TGATCGTTAC CTACCCCCAC	780
TACCATCACT ATCTGGTACT GCAGTGGTTA ACATCCGAGT CCTGGACTTT AATGATAACA ATGGTAGTGA TAGACCATGA CGTCACCAAT TGTAGGCTCA GGACCTGAAA TTACTATTGT	840
GCCCAGTGT TGAGAGAACG ACCATTGCTG TGGACCTAGT AGAGGATGCT CCTCTGGGAT CGGGTCACAA ACTCTCTTCG TGGTAACGAC ACCTGGATCA TCTCCTACGA GGAGACCCCTA	900
ACCTTTGTT GGAGTTACAT GCTACTGACG ATGATGAAGG AGTGAATGGA GAAATTGTTT TGGAAAACAA CCTCAATGTA CGATGACTGC TACTACTTCC TCACTTACCT CTTAACAAA	960
ATGGATTCAAG CACTTGGCA TCTCAAGAGG TACGTCAGCT ATTTAAAATT AACTCCAGAA TACCTAACGTC GTGAAACCGT AGAGTTCTCC ATGCAGTCGA TAAATTAA TTGAGGTCTT	1020

CTGGCAGTGT TACTCTTGA GGCCAAGTTG ATTTTGAGAC CAAGCAGACT TACGARTTG GACCGTCACA ATGAGAACCTT CCGGTTCAAC TAAAACCTCG GTTCGCTCTGA ATGCTTAAAC	1080
AGGTACAAAGC CCAAGATTTG GGCCCCAACC CACTGACTGC TACTTGAAA GTAAGTGTTC TCCATGTTCG GGTTCTAAAC CCGGGGTTGG GTGACTGACG ATGAACATTT CATTGACAAG	1140
ATATACTTGA TGTAAATGAT AATACCCCCAG CCATCACTAT TACCCCTCTG ACTACTGTAA TATATGAACT ACATTTACTA TTATGGGTC GGTAGTGATA ATGGGGAGAC TGATGACATT	1200
ATGCAGGAGT TGCCTATATT CCAGAAACAG CCACAAAGGA GAACTTTATA GCTCTGATCA TACGTCTCA ACGGATATAA GGTCTTGTC GGTGTTCCCT CTTGAAATAT CGAGACTAGT	1260
GCACACTGAA CAGAGCCTCT GGATCTAATG GACAAGTTCG CTGTAACCTT TATGGACATG CGTGATGACT GTCTCGGAGA CCTAGATTAC CTGTTCAAGC GACATGAGAA ATACCTGTAC	1320
AGCACTTTAA ACTACAGCAA GCTTATGAGG ACAGTTACAT GATAGTTACC ACCTCTACTT TCGTGAAATT TGATGTCGTT CGAATACTCC TGTCAATGTA CTATCAATGG TGGAGATGAA	1380
TAGACAGGGA AAACATAGCA GCGTACTCTT TGACAGTAGT TGCAGAAAGAC CTTGGCTTCC ATCTGTCCCT TTTGTATCGT CGCATGAGAA ACTGTCATCA ACGTCTTCTG GAACCGAAGG	1440
CCTCATTGAA GACCAAAAAG TACTACACAG TCAAGGTTAG TGATGAGAAT GACAATGCAC GGAGTAACCTT CTGGTTTTC ATGATGTCG AGTTCCAATC ACTACTCTTA CTGTTACGTG	1500
CTGTATTTTC TAAACCCAG TATGAAGCTT CTATTCTGGA AAATAATGCT CCAGGCTCTT GACATAAAAAG ATTTGGGTC ATACTTCGAA GATAAGACCT TTTATTACGA GGTCCGAGAA	1560
ATATAACTAC AGTGATAGCC AGAGACTCTG ATAGTGATCA AAATGGAAA GTAATTACA TATATTGATG TCACTATCGG TCTCTGAGAC TATCACTAGT TTTACCGTTT CATTAAATGT	1620
GACTTGTGGA TGCAAAAGTG ATGGGCCAGT CACTAACAC ATTTGTTCT CTTGATGCGG CTGAACACCT ACGTTTCAC TACCCGGTCA GTGATTGTTG TAAACAAAGA GAACTACGCC	1680
ACTCTGGAGT ATTGAGAGCT GTTAGGTCTT TAGACTATGA AAAACTAAA CAACTGGATT TGAGACCTCA TAACTCTCGA CAATCCAGAA ATCTGATACT TTTGAATTG GTTGACCTAA	1740
TTGAAATTGA AGCTGCAGAC AATGGGATCC CTCAACTCTC CACTCGCGTT CAACTAAATC AACTTTAACT TCGACGTCTG TTACCCCTAGG GAGTTGAGAG GTGAGCGCAA GTTGATTTAG	1800
TCAGAAATAGT TGATCAAAAT GATAATTGCC CTGTGATAAC TAATCCCTT CTTAATAATG AGTCTTATCA ACTAGTTTA CTATTAACGG GACACTATTG ATTAGGAGAA GAATTATTAC	1860
GCTCGGGTGA AGTTCTGCTT CCCATCAGCG CTCCCTCAAAA CTATTTAGTT TTCCAGCTCA CGAGCCCAC TCAAGACGAA GGGTAGTCGC GAGGAGTTT GATAATCAA AAGGTCGAGT	1920
AAGCCGAGGA TTCAAGATGAA GGGCACAAC CCCAGCTGTT CTATACCATA CTGAGAGATC TTCGGCTCCT AAGTCTACTT CCCGTGTTGA GGGTCGACAA GATATGGTAT GACTCTCTAG	1980
CAAGCAGATT GTTTGCCATT AACAAAGAAA GTGGTGAAGT GTTCCGTAAA AACAAATTAA GTTCGTCTAA CAAACGGTAA TTGTTCTTT CACCACTTCA CAAGGACTTT TTTGTTAATT	2040
ACTCTGACCA TTCAGAGGAC TTGAGCATAG TAGTTGCAGT GTATGACTTG GGAAGACCTT TGAGACTGGT AAGTCTCTG AACTCGTATC ATCAACGTCA CATACTGAAC CCTTCTGGAA	2100
CATTATCCAC CAATGCTACA GTTAAATTCA TCCTCACCGA CTCTTTCCCT TCTAACGTTG GTAATAGGTG GTTACGATGT CAATTTAAGT AGGAGTGGCT GAGAAAAGGA AGATTGCAAC	2160

Fig. 6. (Continuation page 2, SEQ ID NO:6).

AAGTCGTTAT TTTGCAACCA TCTGCAGAAG AGCAGCACCA GATCGATATG TCCATTATAT	2220
TTCAAGCAATA AAACGTTGGT AGACGTCTTC TCGTCGTGGT CTAGCTATAC AGGTAATATA	
TCATTGCAGT GCTGGCTGGT GGTGGTGCCT TGCTACTTTT GGCCATCTTT TTTGTGGCCT	2280
AGTAACGTCA CGACCGACCA CCAACACGAA ACGATGAAAA CCGGTAGAAA AAACACCGGA	
GTACTTGTAAGAAGCT GGTGAATTAA AGCAGGTACG TGAACAAACAC GGAACATGCA	2340
CATGAACATT TTTCTTCGA CCACCTAAAT TCGTCATGG ACTTGTGTTG CTTGTACGT	
ATGAAGAACG CCTGTTAACG ACCCCATCTC CCCAGTCGGT CTCTTCTCT TTGTCAGTCA	2400
TACTTCTTGC GGACAATTG TGTTGAGAG GGGTCAGCCA GAGAAGAAGA AACAGAGTCA	
CTGAGTCATG CCAACTCTCC ATCAAACTG AATCTGAGAA TTGCAGCGTG TCCTCTAAC	2460
GACTCAGTAC GGTTGAGAGG TAGTTATGAC TTAGACTCTT AACGTCGCAC AGGAGATTGG	
AAGAGCAGCA TCAGCAAACA GGCATAAAAGC ACTCCATCTC TGTACCATCT TATCACACAT	2520
TTCTCGTCGT AGTCGTTGT CCGTATTCG TGAGGTAGAG ACATGGTAGA ATAGTGTGTA	
CTGGTTGGCA CCTGGACAAT TGTGCAATGA GCATAAGTGG ACATTCTCAC ATGGGGCACA	2580
GACCAACCGT GGACCTGTTA ACACGTTACT CGTATTCAAC TGTAAAGAGTG TACCCCGTGT	
TTAGTACAAA GGTACAGTGG GCAAAGGAGA TAGTGACTTC AATGACAGTG ACTCTGATAC	2640
AATCATGTTT CCATGTCAAC CGTTCTCT ATCACTGAAG TTACTGTCAC TGAGACTATG	
TAGTGGAGAA TCAGAAAAGA AGAGCATTGA GCAGCCAATG CAGGCACAAG CCAGTGCCTCA	2700
ATCACCTCTT AGTCTTTCT TCTCGTAAC CGTCGGTTAC GTCCGTGTTG GGTACAGGAGT	
ATACACAGAT GAATCAGCAG GGTCCGACA TGCCGATAAC TATTTCAGCC ACCGAATCAA	2760
TATGTGTCTA CTTAGTCGTC CCAAGGCTGT ACGGCTATTG ATAAAGTCGG TGGCTTAGTT	
CAAGGGTCCA GAAAATGGGA ACTGCACATT GCAATATGAA AAGGGCTATA GACTGTCTTA	2820
GTTCCCAAGGT CTTTACCCCT TGACGTGTAA CGTTATACTT TTCCCGATAT CTGACAGAAAT	
CTCTGTAGCT CCTGTATATT ACAATACCTA CCATGCAAGA ATGCCTAACCC TGACACATACC	2880
GAGACATCGA GGACATATAA TGTTATGGAT GGTACGTTCT TACGGATTGG ACAGTGTATGG	
GAACCATACC CTTAGAGACC CTTATTACCA TATCAATAAT CCTGTTGCTA ATCGGATGCA	2940
CTTGGTATGG GAATCTCTGG GAATAATGGT ATAGTTATTA GGACAAACGAT TAGCCTACGT	
GGCGGAATAT GAAAGAGATT TAGTCAACAG AAGTGCAACG TTATCTCCGC AGAGATCGTC	3000
CCGCCTTATA CTTTCTCTAA ATCAGTTGTC TTCACGTTGC AATAGAGGCG TCTCTAGCAG	
TAGCAGATAAC CAAGAATTCA ATTACAGTCC GCAGATATCA AGACAGCTTC ATCCCTCAGA	3060
ATCGTCTATG GTTCTTAAGT TAATGTCAGG CGTCTATAGT TCTGTCGAAG TAGGAAGTCT	
AATTGCTACA ACCTTTAAT CATTAGGCAT GCAAGTGAGA ATGCACAAAG GCAAGTGCTT	3120
TTAACGATGT TGGAAAATTA GTAATCCGTA CGTTCACTCT TACGTGTTTC CGTTCACGAA	
TAGCATGAAA GCTAAATATA TGGAGTCTCC CCTTTCCCTC TGATGGATGG GGGGAGACAC	3180
ATCGTACTTT CGATTTATAT ACCTCAGAGG GGAAAGGGAG ACTACCTACC CCCCTCTGTG	
AGGACAGTGC ATAAATATAC AGCTGCTTTC TATTTGCATT TCACTTGGGA ATTTTTGTT	3240
TCCTGTCACG TATTTATATG TCGACGAAAG ATAAACGTA AGTGAACCCCT TAAAAAAACAA	
TTTTTACAT ATTTATTTT CCTGAATTGA ATGTGACATT GTCCTGTCAC CTAACTAGCA	3300
AAAAAAATGTA TAAATAAAAA GGACTTAAC TACACTGTAA CAGGACAGTG GATTGATCGT	

Fig. 6. (Continuation page 3, SEQ ID NO:6).

ATTAATCCA CAGACCTACA GTCAAATATT TGAGGGCCCC TGAAACAGCA CATCAGTCAG TAATTTAGGT GTCTGGATGT CAGTTATAA ACTCCCGGGG ACTTTGTCGT GTAGTCAGTC	3360
GACCTAAAGT GGCTTTTA CTTTAGCAG CTCCTGGTC TGCCCTCTGT GTTAATCAGC CTGGATTCAGT CCGGAAAAAT GAAAATCGTC GAGGACCCAG ACGGGAGACCA CAATTAGTCG	3420
CCCTGGTCAA GTCTGAGTA GGATCATGGC GTTTTATAT GCATCTCACC TACTTGGAC GGGACCAAGT CAGGACTCAT CCTAGTACCG CAAAAATATA CGTAGAGTGG ATGAAACCTG	3480
GTGATTTACA CATAATAGGA AACGCTTGGT TTCAGTGAAG TCTGTGTGT ATATATTCTG CACTAAATGT GTATTATCCT TTGCGAACCA AAGTCACTTC AGACACAAACA TATATAAGAC	3540
TTATATACAC GCATTTGTG TTTGTGTATA TATTCAGT CCATTCAGAT ATGTGTATAT AATATATGTG CGTAAAACAC AAACACATAT ATAAAGTTCA GGTAAGTCTA TACACATATA	3600
AGTGCAGACC TTGTAATTA AATATTCTGA TACTTTTCC TCAATAATAA TTTAAAT TCACGTCTGG AACATTTAAT TTATAAGACT ATGAAAAGG AGTTATTAT AAATTAA	

Fig. 6. (Continuation page 4, SEQ ID NO:6).

MVCCGPGRML LGWAGLLVLA ALCLLQVPGA QAAACEPVRI PLCKSLPWNM TKMPNHLHHS 60
TQANAILAME QFEGLLGTHC SPDLLFFLCA MYAPICTIDF QHEPIKPCKS VCERARQGCE 120
PILIKYRHSHW PESLACDELP VYDRGVCISP EAIVTADGAD FPMDSSTGHC RGASSERCKC 180
KPVRATQKTY FRNNNYVIR AKVKEVKMKC HDVTAVVEVK EILKASLVNI PRDTVNLYTT 240
SGCLCPPLTV NEEYVIMGYE DEERSRLLL V EGSIAEKWKD RLGKKVKRWD MKLRHLGLGK 300
TDASDSTQNQ KSGRNSNPRP ARS.

Figure 7. Deduced amino acid sequence of mouse FRZB-1 protein. SEQ ID NO:7.

Figure 8. Nucleotide sequence of the full-length mouse FRZB-1 cDNA. SEQ ID NO:8.

AAGCCTGGGA CCATGGTCTG CTGCGGCCCG GGACGGATGC TGCTAGGATG GGCCGGGTTG TTCGGACCT GGTACCAGAC GACGCCGGC CCTGCCTACG ACGATCCTAC CCGGCCAAC	60
CTAGTCCTGG CTGCTCTCTG CCTGCTCCAG GTGCCGGAG CTCAGGCTGC AGCCTGTGAG GATCAGGACC GACGAGAGAC GGACGAGGTC CACGGGCCTC GAGTCCGACG TCGGACACTC	120
CCTGTCCGCA TCCCCTGCTG CAAGTCCCTT CCCTGAAACA TGACCAAGAT GCCCAACCAC GGACAGGCCT AGGGCGACAC GTTCAGGGAA GGGACCTTGT ACTGGTTCTA CGGGTTGGTG	180
CTGCACCCACA GCACCCAGGC TAACGCCATC CTGGCCATGG AACAGTTCGA AGGGCTGCTG GACGTGGTGT CGTGGTCCG ATTGCGGTAG GACCGGTACC TTGTCAAGCT TCCCGACGAC	240
GGCACCCACT GCAGCCCCGA TCTTCTCTTC TTCCCTGTG CAATGTACGC ACCCATTTGC CCGTGGGTGA CGTCGGGCCT AGAAGAGAAG AAGGAGACAC GTTACATGCG TGGGTAAACG	300
ACCATCGACT TCCAGCACGA GCCCATCAAG CCCTGCAAGT CTGTGTGTGA GCGCGCCCGA TGGTAGCTGA AGGTCGTGCT CGGGTAGTTCA GACACACACT CGCGCGGGCT	360
CAGGGCTGCG AGCCCATTCT CATCAAGTAC CGCCACTCGT GGCCGGAAAG CTTGGCCTGC GTCCCCACGC TCGGGTAAGA GTAGTTCATG GCGGTGAGCA CCGGCCTTTC GAACCGGACG	420
GACGAGCTGC CGGTGTACGA CCGCGCGTG TGCATCTCTC CTGAGGCCAT CGTCACCGCG CTGCTCGACG GCCACATGCT GGCGCCGCAC ACGTAGAGAG GACTCCGGTA GCAGTGGCGC	480
GACGGAGCGG ATTTCCAT GGATTCAAGT ACTGGACACT GCAGAGGGGC AAGCAGCGAA CTGCCTCGCC TAAAAGGATA CCTAAGTTCA TGACCTGTGA CGTCTCCCCG TTCGTCGCTT	540
CGTTGCAAAT GTAAGCTGT CAGAGCTACA CAGAAGACCT ATTTCCGGAA CAATTACAAC GCAACGTTA CATTGGACA GTCTCGATGT GTCTCTGGA TAAAGGCCTT GTTAATGTTG	600
TATGTCATCC GGGCTAAAGT TAAAGAGGTA AAGATGAAAT GTCATGATGT GACCGCCGTT ATACAGTAGG CCCGATTCA ATTTCTCCAT TTCTACTTTA CAGTACTACA CTGGCGCAA	660
GTGGAAGTGA AGGAAATTCT AAAGGCATCA CTGGTAAACA TTCCAAGGGAA CACCGTCAAT CACCTTCACT TCCTTTAAGA TTTCCGTAGT GACCATTGT AAGGTTCCCT GTGGCAGTTA	720
CTTTATACCA CCTCTGGCTG CCTCTGTCTT CCACTTACTG TCAATGAGGA ATATGTCATC GAAATATGGT GGAGACCGAC GGAGACAGGA GGTGAATGAC AGTTACTCCT TATAACAGTAG	780
ATGGGCTATG AAGACGAGGA ACGTTCCAGG TTACTCTTGG TAGAAGGCTC TATAGCTGAG TACCCGATAC TTCTGCTCCT TGCAAGGTCC AATGAGAACC ATCTTCCGAG ATATCGACTC	840
AAGTGGAAAGG ATCGGCTTGG TAAGAAAAGTC AAGCGCTGGG ATATGAAACT CCGACACCTT TTCACCTTCC TAGCCGAACC ATTCTTTCAAG TTGCGACCC TATACTTTGA GGCTGTGGAA	900
GGACTGGGTA AAACTGATGC TAGCGATTCC ACTCAGAATC AGAAGTCTGG CAGGAACCT CCTGACCCAT TTTGACTACG ATCGCTAAGG TGAGTCTTAG TCTTCAGACC GTCCTTGAGA	960

699034880 003444000

AATCCCCGGC CAGCACCGAG CTAATCCTG AAATGTAAAA GGCCACACCC ACGGACTCCC	1020
TTAGGGGCCG GTCGTGGTC GATTAGGAC TTTACATTT CCGGTGTGGG TGCCTGAGGG	
TTCTAAGACT GGCCTGGTG GACTAACAAA GGAAAACCGC ACAGTTGTGC TCGTGACCGA	1080
AAGATTCTGA CCGCGACAC CTGATTGTTT CCTTTGGCG TGTCAACACG AGCACTGGCT	
TTGTTTACCG CAGACACCGC GTGGCTACCG AAGTTACTTC CGGTCCCCCTT TCTCCTGCTT	1140
AACAAATGGC GTCTGTGGCG CACCGATGGC TTCAATGAAG GCCAGGGGAA AGAGGACGAA	
CTTAATGGCG TGGGTTAGA TCCTTAATA TGTTATATAT TCTGTTTCAT CAATCACGTG	1200
GAATTACCGC ACCCAAATCT AGGAAATTAT ACAATATATA AGACAAAGTA GTTAGTGCAC	
GGGACTGTTC TTTGCAACC AGAATAGTAA ATAAATATG TTGATGCTAA GGTTCTGTA	1260
CCCTGACAAG AAAACGTTGG TCTTATCATT TAATTATAC AACTACGATT CCAAAGACAT	
CTGGACTCCC TGGTTTAAT TTGGTGTCT GTACCTGAT TGAGAATGCA ATGTTTCATG	1320
GACCTGAGGG ACCCAAATTA AACCACAAGA CATGGACTA ACTCTTACGT TACAAAGTAC	
TAAAGAGAGA ATCCTGGTCA TATCTCAAGA ACTAGATATT GCTGTAAGAC AGCCTCTGCT	1380
ATTTCTCTCT TAGGACCAAGT ATAGAGTTCT TGATCTATAA CGACATTCTG TCGGAGACGA	
GCTGCGCTTA TAGTCTTGTG TTTGTATGCC TTTGTCCATT TCCCTCATGC TGTGAAAGTT	1440
CGACCGAAT ATCAGAACAC AAACATACGG AAACAGGTAA AGGGAGTACG ACACCTTCAA	
ATACATGTTT ATAAAGGTAG AACGGCATTG TGAAATCAGA CACTGCACAA GCAGAGTAGC	1500
TATGTACAAA TATTCCATC TTGCCGTAAA ACTTTAGTCT GTGACGTGTT CGTCTCATCG	
CCAACACCAG GAAGCATTG TGAGGAAACG CCACACAGCA TGACTTATTT TCAAGATTGG	1560
GGTTGTGGTC CTTCGTAAT ACTCCTTGC GGTGTGTCGT ACTGAATAAA AGTTCTAAC	
CAGGCAGCAA AATAAATAGT GTGGGAGCC AAGAAAAGAA TATTTGCCT GGTAAAGGGG	1620
GTCCGTCGTT TTATTTATCA CAACCTCGG TTCTTTCTT ATAAAACGGA CCAATTCCCC	
CACACTGGAA TCAGTAGCCC TTGAGCCATT AACAGCAGTG TTCTTCTGGC AAGTTTTG	1680
GTGTGACCTT AGTCATCGGG AACTCGGTAA TTGTCGTAC AAGAAGACCG TTCAAAAAC	
TTTGTTCATA AATGTATTCA CGAGCATTAG AGATGAACCTT ATAACTAGAC ATCTGTTGTT	1740
AAACAAGTAT TTACATAAGT GCTCGTAATC TCTACTTGAA TATTGATCTG TAGACAACAA	
ATCTCTATAG CTCTGCTTCC TTCTAAATCA AACCCATTGT TGGATGCTCC CTCTCCATTG	1800
TAGAGATATC GAGACGAAGG AAGATTTAGT TTGGGTAACA ACCTACGAGG GAGAGGTAAG	

ATAAAATAAAT TTGGCTTGCT GTATTGGCCA GGAAAAGAAA GTATTAAGT ATGCATGCAT TATTATTTA AACCGAACGA CATAACCGGT CCTTTCTTT CATAATTCA TACGTACGTA	1860
GTGCACCAGG GTGTTATTTA ACAGAGGTAT GTAACTCTAT AAAAGACTAT AATTTACAGG CACGTGGTCC CACAATAAAT TGTCTCCATA CATTGAGATA TTTTCTGATA TTAAATGTCC	1920
ACACGGAAAT GTGCACATTG GTTTACTTTT TTTCTTCCTT TTGCTTTGGG CTTGTGATTT TGTGCCTTTA CACGTGTAAA CAAATGAAAA AAAGAAGGAA AACGAAACCC GAACACTAAA	1980
TGGTTTTTGG TGTGTTTATG TCTGTATTTT GGGGGGTGGG TAGGTTTAAG CCATTGCACA ACCAAAAACC ACACAAATAC AGACATAAAA CCCCCCACCC ATCCAAATTG GGTAAACGTGT	2040
TTCAAGTTGA ACTAGATTAG AGTAGACTAG GCTCATTGGC CTAGACATTA TGATTTGAAT AAGTTCAACT TGATCTAACG TCATCTGATC CGAGTAACCG GATCTGTAAT ACTAAACTTA	2100
TTGTGTTGTT TAATGCTCCA TCAAGATGTC TAATAAAAGG AATATGGTTG TCAACAGAGA AACACAAACAA ATTACGAGGT AGTTCTACAG ATTATTTCC TTATACCAAC AGTTGTCTCT	2160
CGACAAACAAAC AACAAA GCTGTTGTTG TTGTTT	

MVCGSPGGML LLRAGLLALA ALCLLRVPGA RAAACEPVRI PLCKSLPWNM TKMPNHLHHS 60
TQANAILAIE QFEGLLGTHC SPDLLFFLCA MYAPICTIDF QHEPIKPKS VCERARQGCE 120
PILIKYRHSW PENLACEELP VYDRGVCISP EAIVTADGAD FPMDSNGNC RGASSERCKC 180
KPIRATQKTY FRNNYNYVIR AKVKEIKTKC HDVTAVVEVK EILKSSLVNI PRDTVNLYTS 240
SGCLCPPLNV NEEYIIMGYE DEERSRLLL V EGSIAEKWKD RLGKKVKRWD MKLRHLGLSK 300
SDSSNSDSTQ SQKSGRNSNP RQARN.

Figure 9. Deduced amino acid sequence of human FRZB-1 protein. SEQ ID NO:9.

Figure 10. Nucleotide sequence of the full-length human FRZB-1 cDNA. SEQ ID NO:10.

This sequence was assembled from public ESTs from the Genbank database
(accession numbers: H18848, R63748, W38677, W44760, H38379 and N71244).

GGCGGAGCGG GCCTTTGGC GTCCACTGCG CGGCTGCACC CTGCCCATC TGCCGGGATC 60
CCGCCTCGCC CGGAAAACCG CAGGTGACGC GCCGACGTGG GACGGGGTAG ACGGCCCTAG

ATGGTCTGCG GCAGCCCGGG AGGGATGCTG CTGCTGCGGG CGGGCTGCT TGCCCTGGCT 120
TACCAGACGC CGTCGGGCC TCCCTACGAC GACGACGCC GGCCCGACGA ACGGGACCGA

GCTCTCTGCC TGCTCCGGGT GCCCCGGGCT CGGGCTGCAG CCTGTGAGCC CGTCCGCATC 180
CGAGAGACGG ACGAGGCCA CGGGCCCCGA GCCCGACGTC GGACACTCGG GCAGGCGTAG

CCCCTGTGCA AGTCCCTGCC CTGGAACATG ACTAAGATGC CCAACCACCT GCACCACAGC 240
GGGGACACGT TCAGGGACGG GACCTTGTAC TGATTCTACG GGTTGGTGGA CGTGGTGTGCG

ACTCAGGCCA ACGCCATCCT GGCCATCGAG CAGTCGAAG GTCTGCTGGG CACCCACTGC 300
TGAGTCCGGT TGCGGTAGGA CCGGTAGCTC GTCAAGCTTC CAGACGACCC GTGGGTGACG

AGCCCCGATC TGCTCTTCTT CCTCTGTGCC ATGTACGCGC CCATCTGCAC CATTGACTTC 360
TCGGGGCTAG ACGAGAAGAA GGAGACACGG TACATGCGCG GGTAGACGTG GTAAGTGAAG

CAGCACGAGC CCATCAAGCC CTGTAAGTCT GTGTGCGAGC GGGCCCGGCA GGGCTGTGAG 420
GTCGTGCTCG GGTAGTTCGG GACATTAGA CACACGCTCG CCCGGCCGT CCCGACACTC

CCCATACTCA TCAAGTACCG CCACTCGTGG CCGGAGAACCC TGGCCTGCGA GGAGCTGCCA 480
GGGTATGAGT AGTTCATGGC GGTGAGCACC GGCTCTTGG ACCGGACGCT CCTCGACGGT

GTGTACGACA GGGCGGTGTG CATCTCTCCC GAGGCCATCG TTACTGCGGA CGGAGCTGAT 540
CACATGCTGT CCCCCCACAC GTAGAGAGGG CTCCGGTAGC AATGACGCT GCCTCGACTA

TTTCCTATGG ATTCTAGTAA CGGAAACTGT AGAGGGGAA GCAGTGAACG CTGTAAATGT 600
AAAGGATACC TAAGATCATT GCCTTTGACA TCTCCCCGTT CGTCACTTGC GACATTACCA

AAGCCTATTA GAGCTACACA GAAGACCTAT TTCCGGAAACA ATTACAACCA TGTCATTGCG 660
TTCCGGATAAT CTCGATGTGT CTTCTGGATA AAGGCCTTGT TAATGTTGAT ACAGTAAGCC

GCTAAAGTTA AAGAGATAAA GACTAAGTGC CATGATGTGA CTGCAGTAGT GGAGGTGAAG 720
CGATTTCAT TTCTCTATTT CTGATTACAG GTACTACACT GACGTACATCA CCTCCACTTC

GAGATTCTAA AGTCCTCTCT GGTAAACATT CCACGGGACA CTGTCAACCT CTATACCAGC 780
CTCTAAGATT TCAGGAGAGA CCATTTGTTA GGTGCCCTGT GACAGTTGGA GATATGGTCG

TCTGGCTGCC TCTGCCCTCC ACTTAATGTT AATGAGGAAT ATATCATCAT GGGCTATGAA 840
AGACCGACGG AGACGGGAGG TGAATTACAA TTACTCCTTA TATAGTAGTA CCCGATACTT

GATGAGGAAC GTTCCAGATT ACTCTTGGTG GAAGGCTCTA TAGCTGAGAA GTGGAAGGAT CTACTCCTTG CAAGGTCTAA TGAGAACAC CTTCCGAGAT ATCGACTCTT CACCTTCCTA	900
CGACTCGGTA AAAAAGTTAA GCGCTGGGAT ATGAAGCTTC GTCATCTTGG ACTCAGTAAA GCTGAGCCAT TTTTCATT CGCGACCCA TACTTCGAAG CAGTAGAAC TGAGTCATTT	960
AGTGATTCTA GCAATAGTGA TTCCACTCAG AGTCAGAAGT CTGGCAGGAA CTCGAACCCC TCACTAAGAT CGTTATCACT AAGGTGAGTC TCAGTCTTCA GACCGTCCTT GAGCTTGGGG	1020
CGGCAAGCAC GCAACTAAAT CCCGAAATAC AAAAAGTAAC ACAGTGGACT TCCTATTAAG GCCGTTCGTG CGTTGATTAA GGGCTTATG TTTTCATTG TGTCACCTGA AGGATAATTC	1080
ACTTACTTGC ATTGCTGGAC TAGCAAAGGA AAATTGCACT ATTGCACATC ATATTCTATT TGAATGAACG TAACGACCTG ATCGTTCTT TTTAACGTGA TAACGTGTAG TATAAGATAAA	1140
GTTTACTATA AAAATCATGT GATAACTGAT TATTACTTCT GTTCTCTTT TGGTTCTGC CAAATGATAT TTTAGTACA CTATTGACTA ATAATGAAGA CAAAGAGAAA ACCAAAGACG	1200
TTCTCTCTTC TCTCAACCCC TTTGTAATGG TTTGGGGCA GACTCTTAAG TATATTGTGA AAGAGAGAAG AGAGTGGGG AAACATTACC AAACCCCCGT CTGAGAATTC ATATAACACT	1260
GTTTCTATT TCACTAATCA TGAGAAAAAC TGTTCTTTG CAATAATAAT AAATTAAACA CAAAGATAA AGTGATTAGT ACTCTTTTG ACAAGAAAAC GTTATTATTA TTTAATTGT	1320
TGCTGTTACC AGAGCCTCTT TGCTGAGTCT CCAGATGTTA ATTTACTTTC TGCAACCCAA ACGACAATGG TCTCGGAGAA ACGACTCAGA GGTCTACAAT TAAATGAAAG ACGTGGGTT	1380
TTGGGAATGC AATATTGGAT GAAAAGAGAG GTTCTGGTA TTCACAGAAA GCTAGATATG AACCTTACG TTATAACCTA CTTTCTCTC CAAAGACCAT AAGTGTCTT CGATCTATAC	1440
CCTTAAAACA TACTCTGCCG ATCTAATTAC AGCCTTATTT TTGTATGCCCT TTTGGCATT GGAATTGT ATGAGACGGC TAGATTAATG TCGGAATAAA AACATACGGA AAACCGTAA	1500
CTCCTCATGC TTAGAAAGTT CCAAATGTTT ATAAAGGTAA AATGGCAGTT TGAAGTCAAA GAGGAGTACG AATCTTCAA GGTTTACAAA TATTCCATT TTACCGTCAA ACTTCAGTTT	1560
TGTCACATAG GCAAAGCAAT CAAGCACCAAG GAAGTGTGTTA TGAGGAAACA ACACCCAAGA ACAGTGTATC CGTTCTGTTA GTTCGTGGTC CTTCACAAAT ACTCCTTGT TGTGGTTCT	1620
TGAATTATTT TTGAGACTGT CAGGAAGTAA AATAAATAGG AGCTTAAGAA AGAACATTTT ACTTAATAAA AACTCTGACA GTCCTTCATT TTATTTATCC TCGAATTCTT TCTTGTAAAA	1680
GCCTGATTGA GAAGCACAAAC TGAAACCAGT AGCCGCTGGG GTGTTAATGG TAGCATTCTT CGGACTAACT CTTCGTGTG ACTTTGGTCA TCGGCGACCC CACAATTACC ATCGTAAGAA	1740
CTTTGGCAA TACATTGAT TTGTTCATGA ATATATTAAT CAGCATTAGA GAAATGAATT GAAAACCGTT ATGTAACACT AACAAAGTACT TATATAATTA GTCGTAATCT CTTTACTTAA	1800
ATAACTAGAC ATCTGCTGTT ATCACCATAG TTTGTTTAA TTTGCTTCTT TTTAAATAAA TATTGATCTG TAGACGACAA TAGTGGTATC AAAACAAATT AACGAAGGA AAATTTATTT	1860
CCCATTGGTG AAAGTCAAAA AAAAAAAA AAA GGGTAAACCAC TTTCAGTTTT TTTTTTTTT TTT	